

# Master Thesis

Contrast-Agnostic Groupwise Registration by  
Robust PCA for Quantitative Cardiac MRI

Xinqi Li

Delft University of Technology

# Master Thesis

## Contrast-Agnostic Groupwise Registration by Robust PCA for Quantitative Cardiac MRI

by

Xinqi Li

Student Name	Student Number
--------------	----------------

Xinqi Li	5478073
----------	---------

Thesis advisor: Jan van Gemert

Co-supervisor: Qian Tao

Project Duration: November 2022 - June 2023

Institution: Delft University of Technology

Faculty: Faculty of Electrical Engineering, Mathematics    Computer Science

Research Group: Pattern Recognition and Bioinformatics

# Preface

This thesis report summarises my thesis work, which is the final part of the Master of Science in Computer Science at the Delft University of Technology. Two-year master's study passed so fast and I'm really happy to have the chance to continue my research about medical image analysis as my thesis topic.

I really appreciate the help from my co-supervisor, Qian Tao, who guided me throughout the project. I learned a lot about how to find the problem during the research and how to properly analyse and solve them. Besides, thanks my thesis supervisor Jan van Gemert for always insightful suggestions.

Finally, I would like to thank my parents and friends for all your support and help!

*Xinqi Li*  
*Delft, June 2023*

# Contents

<b>Preface</b>	<b>i</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Scientific Paper</b>	<b>3</b>
<b>3 Quantitative Cardiac MRI</b>	<b>13</b>
3.1 Quantitative Cardiac MRI . . . . .	13
3.1.1 T1 mapping . . . . .	14
3.1.2 Modified Look-Locker Pulse Sequence . . . . .	14
<b>4 Deformable Image Registration</b>	<b>16</b>
4.1 DIR frameworks: Pairwise and groupwise . . . . .	16
4.2 DIR frameworks: Conventional and Deep Learning-based . . . . .	17
4.3 Similarity Criterion . . . . .	18
4.3.1 Mean Square Error . . . . .	18
4.3.2 Cross Correlation . . . . .	18
4.3.3 Mutual Information . . . . .	18
4.3.4 Total Correlation . . . . .	19
4.4 B-spline Transform Model . . . . .	20
4.5 Robust Principal Component Analysis . . . . .	20
<b>5 Numerical Phantoms Synthesis</b>	<b>22</b>
<b>6 Additional Experiments</b>	<b>24</b>
6.1 Round Selection . . . . .	24
6.2 Comparison of different similarity metrics . . . . .	25
6.2.1 SD Difference . . . . .	25
6.2.2 Examination of potential artefacts . . . . .	25
<b>References</b>	<b>27</b>

# 1

## Introduction

Magnetic resonance imaging (MRI) is a radiation-free medical imaging technique that provides excellent soft tissue contrast for diagnostic purpose. In cardiovascular disease diagnosis, quantitative cardiac magnetic resonance (CMR), particularly T1 mapping, plays a vital role. T1 mapping is derived from fitting exponential recovery curves of multiple T1-weighted images on a pixel-by-pixel basis. However, the presence of uncorrected cardiac and respiratory motion artifacts in the baseline images poses a challenge in generating precise, reliable, and high-quality quantitative mappings. To enhance the quality of quantitative Cardiac MRI, motion correction using deformable image registration is a crucial step in the post-processing pipeline.

Deformable image registration aims to establish voxel-to-voxel correspondences among a series of acquired images. Traditional methods, such as subject-specific approaches, typically optimize a cost function (e.g., an energy equation) iteratively for each subject, yielding promising results but requiring considerable time. Recently, deep learning-based methods have gained popularity in medical image registration due to their robust performance and fast inference process. However, several challenges exist in building a deep-learning deformable image registration method for quantitative cardiac MRI:

- The deep-learning-based registration methods are usually implemented in a pairwise manner, posing difficulties in selecting a *best* fixed image and propagating registration errors, especially for cardiac MRI. An alternative approach to register quantitative Cardiac MRI is groupwise image registration, which registers all images simultaneously. The groupwise methods register a series of images by optimizing a global registration metric, such as the normalized cross-correlation (NCC) and mutual information (NMI).
- Quantitative cardiac MRI involves rapid changes in image contrast and intensity across a series of images, which makes it challenging to design a consistently reliable optimization metric. Widely-used metrics like NCC and NMI can still be sensitive to contrast changes and may fail.
- To evaluate the quality of generated T1 maps, a common method is to measure the estimated standard deviation (SD) error, which is only an indication of the curve fitting error. However, real datasets often lack gold standard T1 maps, making it necessary to validate the method using synthetic datasets with known *true* T1 maps.

In this work, we proposed a modular, groupwise image registration framework based on robust principle component analysis (rPCA) for fast, robust motion correction in contrast-agnostic sequences. We evaluate the proposed method on both pre-contrast and post-contrast cardiac MRI images and further validate it using numerical phantoms with gold standard T1 maps.

The rest of this thesis report is structured as follows: Chapter 2 presents the scientific article describing the proposed method in this thesis, Chapter 3 provides essential technical details of quantitative cardiac MRI, Chapter 4 introduces basic information regarding deformable image registration, Chapter 5 illustrates the fundamental information of phantom synthesis and Chapter 6 includes additional experiment results. The github repository for the whole project can be found here: <https://github.com/lixinqi98/Master-thesis/tree/master-thesis>

2

Scientific Paper

# Contrast-Agnostic Groupwise Registration by Robust PCA for Quantitative Cardiac MRI

Xinqi Li

Delft University of Technology  
Delft, The Netherlands

X.Li-63@student.tudelft.nl

## Abstract

*Quantitative cardiac MRI is an increasingly important diagnostic tool for cardiovascular diseases. Yet, it is essential to have correct image registration for good accuracy and precision of quantitative mapping. Registering all baseline images from a quantitative cardiac MRI sequence, however, is nontrivial because the patient is moving, leading to simultaneous changes in motion, intensity, and contrast. The changes in image contrast, in particular, make it challenging to design a reliable registration metric for optimization.*

*In this paper, we propose a novel approach based on robust principle component analysis (rPCA) that decomposes quantitative cardiac MRI into low-rank and sparse components, in combination with a groupwise CNN-based registration backbone. The proposed framework aims for fast, robust motion correction for contrast-agnostic sequences, which benefits registration. We evaluated our proposed method on cardiac T1 mapping sequences, both pre-contrast and post-contrast. Additionally, we synthesize the numerical phantoms with gold standard to test the performance. Our experiments showed that our method effectively improved registration performance over baseline methods without rPCA, and reduced quantitative mapping error in both in-domain and out-of-domain MRI sequences. The proposed rPCA framework is generic and can be easily incorporated into existing registration methods and other clinical applications.*

## 1. Introduction

Quantitative cardiac MRI, such as T1 mapping [26], is an increasingly important cardiovascular imaging modality to examine cardiac tissue [11]. However, the quality of quantitative mapping is negatively affected by respiratory and cardiac motion during the MR acquisition procedure [32]. Such motion leads to misalignment of MR sig-

nals across baseline images and results in deteriorated accuracy and precision of quantitative mapping [21]. To improve the quality of quantitative cardiac MRI, motion correction by deformable image registration is an essential part of the post-processing pipeline [2, 9, 25].

Conventionally, the deformable image registration is implemented in a pairwise fashion where two images are registered, with one designated as fixed and one moving. However, for quantitative cardiac MRI, the number of images is highly variable (ranging from 3 to  $> 20$ ), depending on the specific sequence. This makes pairwise registration non-intuitive, as the “best” fixed image is hard to define. Moreover, registration error easily propagates across the baseline images given all pairwise registration steps are independently performed. An alternative approach called *groupwise image registration*, which instead registers all images simultaneously, has gained popularity for quantitative MRI [15, 16, 20, 30]. Groupwise registration promises improved robustness across a sequence of images by optimizing a global registration metric, such as normalized cross-correlation (NCC) and mutual information (NMI). However, an under-studied problem is that degenerated solutions, in the format of ghosting artifacts or pixel collapse [12], may occur during groupwise registration and severely undermine the clinical application. In this paper, we will follow the groupwise registration principle and investigate the susceptibility of NCC and NMI to artifacts.

An additional challenge related to quantitative cardiac MRI is that the change in image contrast and intensity can vary drastically across baseline images with different signal models, and agnostic to the image registration pipeline [32]. Even with the same signal model, the contrast in individual baseline images still depends on the exact scheme of acquisition, which differs among vendors and centers. This makes it difficult to design a consistently reliable registration metric for optimization. Conventional registration metrics, such as NCC and NMI, can still be sensitive to agnostic contrast changes and fail [5, 12, 22]. Therefore, it's of great interest to find a robust registration metric that can work

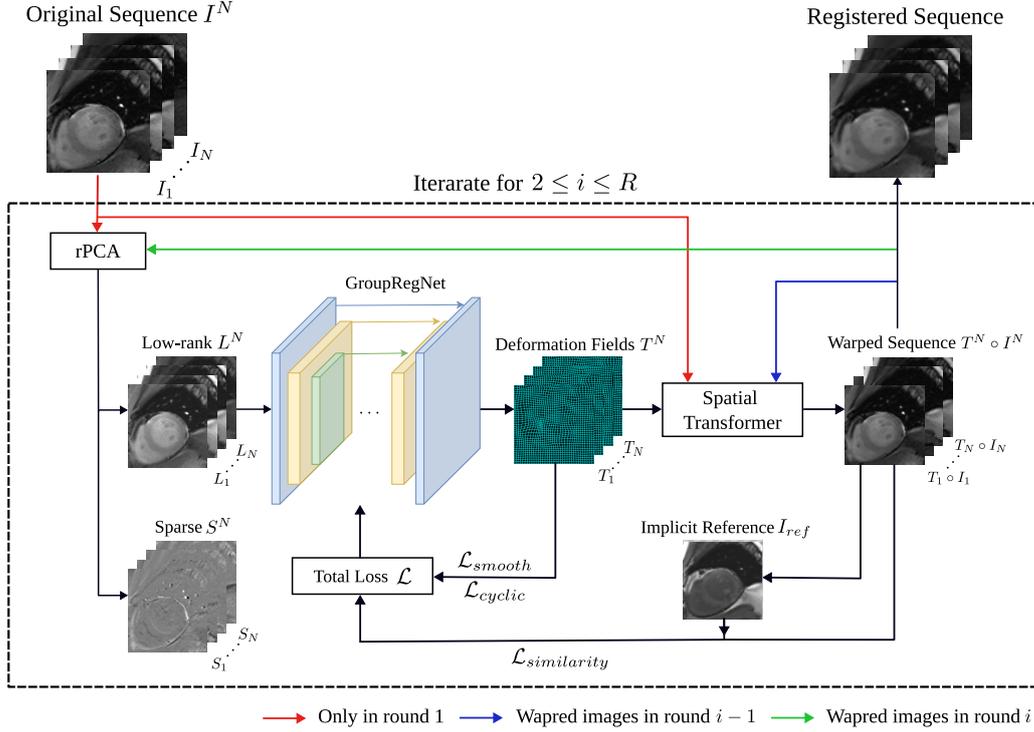


Figure 1. Overview of the proposed framework for contrast-agnostic registration. The dotted rectangle denotes the iterative registration pipeline that progressively corrects motion from round 2 to maximal round  $R$ .

despite agnostic contrast changes.

Another one long-standing problem in evaluating the image registration is that there is no perfect T1 mapping for real dataset, as all quantitative cardiac MRI are a estimation of the perfect result. Due to the lack of the gold standard, the existing measurement can only be treated as indicator of the performance. Recently, physics-based digital phantom simulations are proposed as a powerful tool to evaluate and test the medical image analysis. 4D eXtended CArdiac Torso(XCAT) [29] is developed to provide a realistic and flexible anatomical model of the human, which captures phantoms of varying ages from newborn to adult. MRXCAT2.0 [6] phantom framework further enables the demonstration of cine [23] and first-pass [31] myocardial perfusion imaging with adjustable tissue percentage, dynamic contrast, signal models, multiple receiver coils and noise.

In this paper, we aim to tackle the agnostic contrast change in quantitative cardiac MRI by designing a generic registration framework, which integrates robust PCA (rPCA) [7] with state-of-the-art image registration modules. Our rationale is as follows: firstly, the signal model, which is typically governed by physics principles, has a limited degree of freedom [10, 26], and underlies the *low-rank* component of rPCA. Secondly, the motion of

quantitative cardiac MRI is peculiar in the sense that it is only concentrated around the heart, induced by non-ideal breath-hold and heart rate variability, while the background, e.g., rib cage and lung, stay largely static. The motion field is therefore sparse, corresponding to the *sparsity* component of rPCA. In addition, by taking all baseline images into account, rPCA naturally fits in the groupwise registration regime.

Groupwise image registration can be divided into two paradigms: classical iterative optimization methods that are relatively slow [15, 18–20, 24, 30] and deep-learning-based methods that promise fast inference [1, 8, 13, 17, 33]. In this paper, we propose to integrate rPCA with the state-of-the-art deep-learning groupwise registration backbone [33] for fast, reliable motion correction of quantitative cardiac MRI. Our main contributions are:

1. We propose a novel groupwise image registration framework, which is, to the best of our knowledge, the first attempt to utilize rPCA in groupwise registration with a deep learning backbone.
2. We offer a modular framework, which can be incorporated into any existing registration methods, either classical optimization or modern deep learning methods.

3. We evaluated and demonstrated the generalizability of our contrast-agnostic method on out-of-domain quantitative MRI sequences.
4. We evaluated our contrast-agnostic method with synthetic numerical phantoms thoroughly.

In addition, we further investigated the fitness of two popular metrics, NCC and NMI, for groupwise registration. With our experiments, we showed that NCC could lead to potential artifacts in quantitative mapping, which are detrimental to the clinical application and often overlooked.

## 2. Methods

### 2.1. Problem Formulation

Given a sequence of images  $I^N = \{I_i \in \mathbb{R}^{H \times W} | i = 1, \dots, N\}$ , the goal of groupwise registration is to align all  $I_i$  into one common coordinate system by obtaining a set of deformation fields  $T^N = \{T_i \in \mathbb{R}^{2 \times H \times W} | i = 1, \dots, N\}$ . In our framework, an implicit reference  $I_{ref} = \frac{1}{N} \sum_{n=1}^N (T_n \circ I_n)$ , is generated and help optimize the neural network. Therefore, each  $T_i$  should align anatomical structures in quantitative cardiac MRI, between  $I_i$  and an implicit reference  $I_{ref}$ .

We strive for robust motion correction in a sequence of varying contrasts using rPCA iteratively in our framework. The image sequence  $I^N$  is first decomposed with rPCA to low-rank matrix  $L^N$  and sparse matrix  $S^N$ . Then  $L^N$  is used for training the deep learning backbone and generating deformation field  $T^N$ . Then  $T^N$  is applied to raw images  $I^N$  to get warped images  $T^N \circ I^N$ . The warped images are processed with rPCA again and the previous processing is repeated until the maximal iteration number is reached. An overview of our proposed method is shown in Figure 1.

### 2.2. Robust Principal Component Analysis

Robust principal component analysis (rPCA) [7], as its name suggests, is a more robust matrix decomposition compared to PCA: For a given data matrix  $M$ , where in our case,  $M$  is the matrix of vectorized grouped images  $I^N$ , the rPCA decomposes  $M \in \mathbb{R}^{m \times n}$  into the sum of a low-rank matrix  $L$  and a sparse matrix  $S$  via solving the following optimization problem:

$$\text{minimize } \|L\|_* + \lambda \|S\|_1, \text{ subject to } L + S = M, \quad (1)$$

where  $\|\cdot\|_*$  denotes the nuclear norm,  $\|\cdot\|_1$  denotes the  $l_1$  norm, and  $\lambda$  is a hyperparameter used for the trade-off between the two components, which is often set by default as  $\lambda = 1/\sqrt{\max(m, n)}$ . Such optimization problems can be solved by well-established algorithms, such as proximal (stochastic) gradient descent methods [14].

For quantitative cardiac MRI, the motion has a low degree of freedom and thus can be captured by a low-rank

component  $L$ , while the signal changes can be captured by the sparse part  $S$  as it is mostly concentrated around the heart boundary. An illustration of rPCA on pre-contrast cardiac MRI is shown in Figure 2.

### 2.3. Loss Functions

The optimization problem for finding the deformable mapping  $T^N$  can be formulated as follows:

$$T^N = \arg \min_{T^N} \mathcal{L}_{\text{similarity}} + \lambda_0 \mathcal{L}_{\text{smooth}} + \lambda_1 \mathcal{L}_{\text{cyclic}}, \quad (2)$$

where  $\mathcal{L}_{\text{similarity}}$ ,  $\mathcal{L}_{\text{smooth}}$ , and  $\mathcal{L}_{\text{cyclic}}$  denote similarity function, smoothness regularization, and cyclic consistency, respectively, with trade-off parameters  $\lambda_0$  and  $\lambda_1$ .

**Similarity Functions:** We employed normalized mutual information (NMI) to measure the similarity between the input images  $I^N$  to the warped images  $T^N \circ I^N$ , which is robust when a linear relation between image intensities does not exist [12]. The NMI between two images is defined as:

$$NMI(I_1, I_2) = \frac{2MI(I_1, I_2)}{H(I_1) + H(I_2)}, \quad (3)$$

where  $MI(I_1, I_2)$  denotes the mutual information between  $I_1$  and  $I_2$ ,  $H(I_1)$  is the entropy of image  $I_1$ , and  $H(I_2)$  for image  $I_2$ , respectively. For groupwise registration, the similarity loss  $\mathcal{L}_{\text{similarity}}$  is then defined as:

$$\mathcal{L}_{\text{similarity}} = -\frac{1}{N} \sum_{n=1}^N NMI(T_n \circ I_n, I_{ref}). \quad (4)$$

Another similarity loss is also considered and discussed, which is the local normalized cross-correlation (NCC) [3], defined as:

$$NCC(I_1, I_2) = \frac{1}{H \times W} \sum_{i,j \in H,W} \frac{\sum_{x \in \Omega} (I_1(x) - \bar{I}_1(i, j))(I_2(x) - \bar{I}_2(i, j))}{\sqrt{\hat{I}_1(i, j)\hat{I}_2(i, j)}},$$

where  $\Omega$  indicates the neighborhood voxels around the voxel at position  $(i, j)$  and  $\bar{I}_1(i, j)$  and  $\hat{I}_1(i, j)$  denote the local mean and variance.

**Smoothness Regularization:** The smoothness of the deformation field is regularized through B-spline registration [28]. We adopted B-spline because it can prevent the image from folding and inherently lead to smooth deforma-

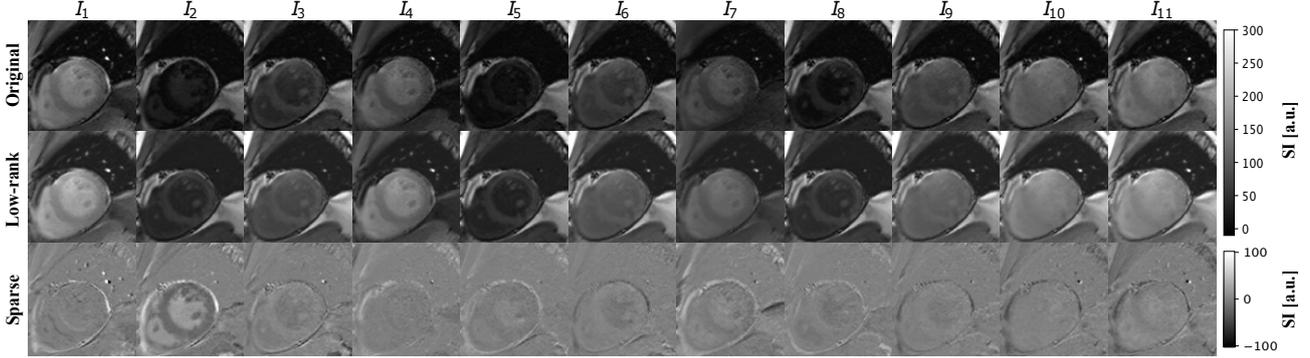


Figure 2. Decomposition of pre-contrast MOLLI cardiac time-series images using rPCA. Each MOLLI sequence consists of 11 pre-contrast time frames in our setting. SI denotes signal intensity. The intensity inconsistency of the sequence is mitigated as shown in low-rank matrix, for example, image  $I_4$ ,  $I_5$  and  $I_8$ .

tion fields:

$$\mathcal{L}_{\text{smooth}} = \frac{1}{H \times W} \sum_{n=1}^N \int_0^H \int_0^W \left[ \left( \frac{\partial^2 \hat{T}_n}{\partial x^2} \right)^2 + \left( \frac{\partial^2 \hat{T}_n}{\partial y^2} \right)^2 + 2 \left( \frac{\partial^2 \hat{T}_n}{\partial xy} \right)^2 \right] dx dy,$$

where  $\hat{T}_n = T_n + \sum_{l=0}^k \sum_{m=0}^k B_l(u) B_m(v) \phi_{i+l, j+m}$ , and where  $B_l$  is the  $l$ -th B-spline basis function,  $k$  is the order of B-spline, and  $\phi_{i,j}$  denotes the control points with uniform space across the image. The control points of the B-spline affect the surrounding deformation fields only the B-spline basis functions.

**Cyclic Consistency:** For groupwise registration, the cyclic consistent regularization keeps the estimated implicit reference at the center of all input images in the manifold by minimizing the deformation field to the implicit reference [33]:

$$\mathcal{L}_{\text{cyclic}}(T^N) = \sqrt{\frac{1}{2(H \times W)} \sum_{i,j \in H,W} \left( \sum_n T_n(i,j) \right)^2}, \quad (5)$$

where  $T_n(i,j)$  denotes the value of  $T_n$  at coordinate  $(i,j)$ . This term prevents the degenerated solution where textures in all images collapse.

## 2.4. CNN-based Neural Network Architecture

The convolution neural network architecture follows that of the VoxelMorph [4], and GroupRegNet [33], based on the UNet [27] architecture consisting of encoding and decoding layers with skip connection. Both the encoder and decoder use convolutional blocks consisting of a 2D convolution and a Leaky ReLU layer. The encoder captures the hierarchical features of the input images with multiple convolution

blocks. The number of decoder layers was controlled by B-spline degree  $k$  [27], the number of control points for dense bspline transform. This enables the coarse-to-fine representation of the two-channel deformation field. The final deformation field is computed by B-spline free form deformation (FFD) transformation model [28] based on the decoder output.

## 2.5. Evaluation Methods

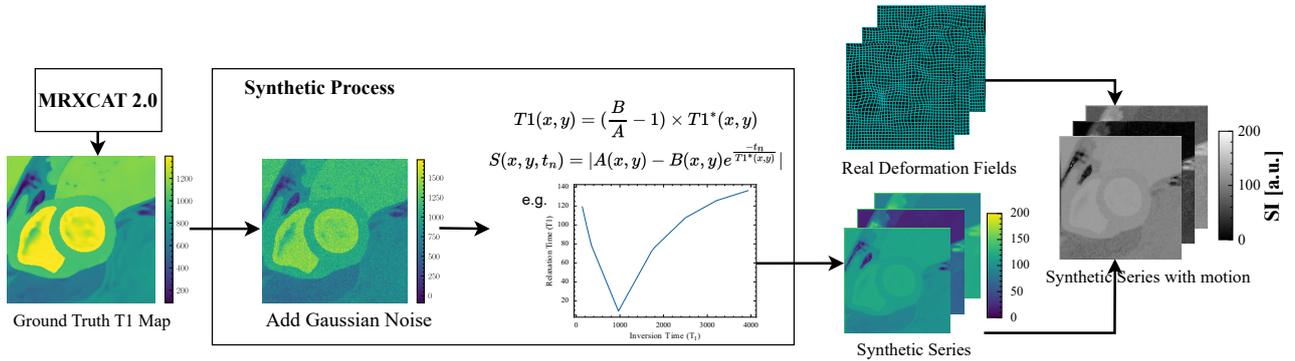
**Quantitative cardiac MRI:** In this paper, we used myocardial T1 mapping, one of the most widely used quantitative mapping modalities in clinic [26]. T1 mapping follows a three-parameter model, expressed by

$$y(T_I) = A - B e^{-T_I/T_1^*}, \quad (6)$$

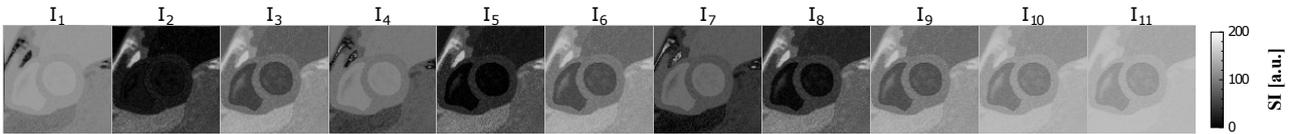
where  $y$  denotes the signal intensity,  $T_I$  denotes the inversion time for acquisition of each baseline image, and  $A$ ,  $B$ , and  $T_1^*$  are parameters to be estimated. Since motion correction leads to a better fitting of this MR physics model at each pixel, here we measure the performance through the T1 mapping within the ROI (myocardium and left ventricle) and the standard deviation (SD) error [21] as an indication of the fitting error. A lower SD error indicates better motion correction. We used both the native (pre-contrast) T1 mapping and post-contrast T1 mapping sequences (after Gadolinium administration). To test the generalizability of our framework, we trained our NN exclusively on pre-contrast T1 mapping, while testing it on both pre-contrast (in-domain) and post-contrast (out-of-domain) sequences.

**Dissimilarity metrics  $\mathcal{D}_{PCA}$ :** We evaluate the warped images using  $\mathcal{D}_{PCA}$ , the ratio of the top- $K$  eigenvalues to the sum of eigenvalues of the correlation matrix [20]. The higher the ratio, the better the performance of registration.

**Baseline methods:** We compared our proposed framework with two methods: (1) the conventional PCA-based groupwise registration method [20], and (2) the groupwise registration method [33] without rPCA, denoted by



(a) Overview of the numerical digital phantom's synthesis. The synthetic data is generated based on the ground truth T1 map and real deformation fields according to quantitative mapping modalities. The plot in synthetic process demonstrates an generating example with user-defined parameters.



(b) One random sample of numerical phantoms. The contrast changes and the cardiac motion follow the real influence of MOLLI sequences.

Figure 3. Synthesis of realistic numerical phantoms. Fig. 3a illustrates the experimental workflow of data synthesis. Figure 3b demonstrates an example of synthetic data.

*GroupRegNet\**. The conventional PCA method minimizes a PCA-based cost function. The groupwise registration method follows [33], but using NMI as the optimization metric. We also performed experiments on *GroupRegNet\** using the NCC metric as in the original work and compared the results with NMI.

**Synthesis of realistic numerical phantoms:** Previous approaches of CMR registration usually measure the performance without the *real* ground truth, relying instead on metrics such as the standard deviation (SD) error and dissimilarity metric  $\mathcal{D}_{PCA}$  to approximately indicate the performance. In our work, we address this limitation by generating an auxiliary validation dataset, including realistic numerical phantoms and corresponding *real* T1 maps  $T1_{real}$ . The synthesis pipeline is illustrated in Figure 3a. To generate the ground truth T1 maps, we utilize MRXCAT2.0 [6], which provides known and detailed functional ground truth of LV morphology and function. In the synthetic process, we add the Gaussian noise to the map and subsequently generate synthetic data following the MOLLI contrast agent dynamics as follows:

$$T1(x, y) = \left(\frac{B}{A} - 1\right) \times T1^*(x, y) \quad (7)$$

$$S(x, y, t_n) = \left|A(x, y) - B(x, y) \times e^{\frac{-t_n}{T1^*(x, y)}}\right|, \quad (8)$$

where  $A(x, y)$ ,  $B(x, y)$  are user-defined parameters and  $TI$  inversion times can be obtained from the real data. The de-

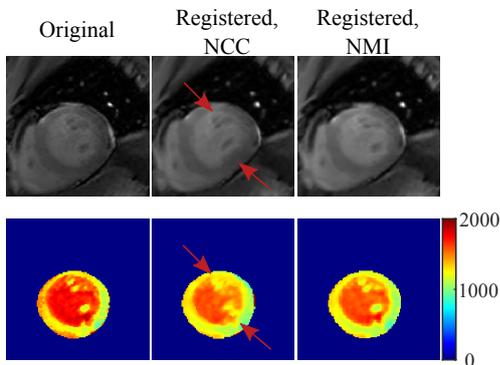
formation fields from the real data is applied on the synthetic series with contrast changes to generate the synthetic series with motion and contrast changes. For example, Figure 3b shows one random sample of the realistic numerical phantoms. The performance evaluation of image registration on numerical phantoms extends beyond previous metrics and includes the calculation of the Root Mean Square Error (RMSE) over the region of interest (ROI) between the real T1 maps  $T1_{real}$  and the estimated T1 maps.

### 3. Experiments and Results

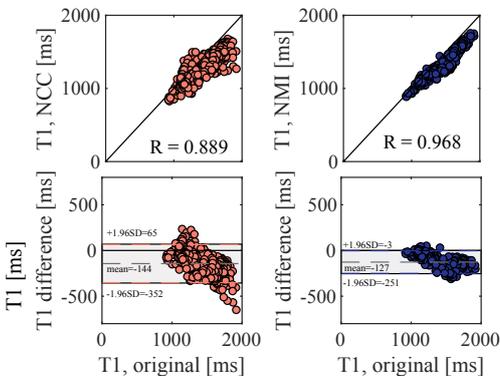
**Dataset:** We used a cardiac MRI dataset including 48 subjects' pre-contrast and post-contrast MOLLI sequences (Philips 3.0T). Each subject had 1 to 3 different slices acquired at the base, mid-ventricular, and apex levels. In total 120 pre-contrast and 120 post-contrast MOLLI sequences were included. All images were resampled to a  $224 \times 224 \times 11$  grid with  $1 \text{ mm}^3$  isotropic resolution and then cropped to  $112 \times 112 \times 11$  at the center. The training comprised 100 random images from only the pre-contrast MOLLI sequences. The rest 20 pre-contrast MOLLI sequences and their corresponding post-contrast sequences, in total 40, formed the test set. We note here that the pre-contrast sequences are the *in-domain* test data, while the post-contrast sequences are the *out-of-domain* test data, given their contrast changes follow a vastly different pattern.

Table 1. Experiment results on T1 mapping. We compare T1 SD and  $\mathcal{D}_{PCA}(K = 1)$  before and after registration and report the percentage of change. The SD diff measures the T1 mapping quality within the ROI (myocardium and left ventricle). Higher values indicate better performance of registration and quantitative mapping. Our method (w/ rPCA) outperforms the GroupRegNet\* on both pre-contrast and post-contrast data according to the SD diff and  $\mathcal{D}_{PCA}$ .

Method	Modality	SD diff	$\mathcal{D}_{PCA}$	Time (s)
ConventionalPCA	Pre-Gd (in domain)	<u>0.33</u>	-0.03	$\approx 600$
GroupRegNet*	Pre-Gd (in domain)	0.33	0.64	1.28
Ours (w/ rPCA)	Pre-Gd (in domain)	<b>0.34</b>	<b>0.76</b>	7.11
ConventionalPCA	Post-Gd (out of domain)	<b>0.26</b>	0.07	$\approx 600$
GroupRegNet*	Post-Gd (out of domain)	0.11	-0.31	1.28
Ours (w/ rPCA)	Post-Gd (out of domain)	<u>0.18</u>	<b>0.52</b>	7.11



(a) An exemplar case of the original and registered images using NCC (middle) and NMI (right). The second row shows the resulting T1 map.



(b) The correlation and Bland-Altman plot of the quantitative T1 estimation within the myocardium ROI.

Figure 4. Comparison between NCC and NMI on T1 mapping. Red arrows in Fig. 4a point to the potential deformation artifact of the NCC-guided registration, in the form of implausible anatomical deformation and biased T1 estimation Fig. 4b.

**Implementation Details:** Robust PCA was implemented with the GoDec algorithm [34]. In each round, the rank of  $L$  was set to be half of the sequence length, which was 5 in our case. Empirically we applied a de-

coder with 4 layers. In this case, the decoder included 2 convolution blocks and the output deformation field was  $31 \times 31 \times 11 \times 2$ . The final deformation field was transformed to  $112 \times 112 \times 11 \times 2$  using B-spline FFD, where the number of control points for dense bspline transform was set 4 and the number of smooth vector field (SVF) integration steps was set 7. The smooth regulation’s weight  $\lambda_0$  is set to 0.001 and cyclic regulation’s weight  $\lambda_1$  is set to 0.01 empirically. We chose to run 3 rounds for our rPCA based method because the performance didn’t change a lot after 3 rounds to save the running time.

**Choice of Similarity Functions:** Two similarity functions, NCC and NMI, are evaluated. We compared the metrics by training on the same training data and checked the result after registration and parametric fitting. We observed that the NCC loss led to undesirable deformation of the ROI, as well as altered distribution of the T1 values, as illustrated in Figure 4. As suggested in [12], NCC might favor homogeneous distribution of pixel intensities and lead to over-smooth myocardium textures that fail the purpose of quantitative mapping. In comparison, NMI better maintained the shape and texture of the ROI.

**Results of Real Data:** The quantitative results of registration and quantitative mapping are shown in Table 1. Note that to demonstrate the generalizability of the learned model, we train the model **only** on pre-contrast data (denoted as Pre-Gd) and tested on both pre- and post-contrast data. Our method performs best on pre-contrast datasets according to the percentage of SD diff within the ROI and outperforms the GroupRegNet\* on post-contrast data. The conventional PCA gives better performance on post-contrast data due to the subject-based optimization nature. However, it takes around 10 minutes for each subject, which is much slower compared to our method, with an average inference time of 7.1 s per sequence with 11 images. One random result of T1 mapping with different methods can be seen in Figure 5.

**Numerical Phantoms Synthesis and Results:** Given the limited availability of public data from MRXCAT2.0,

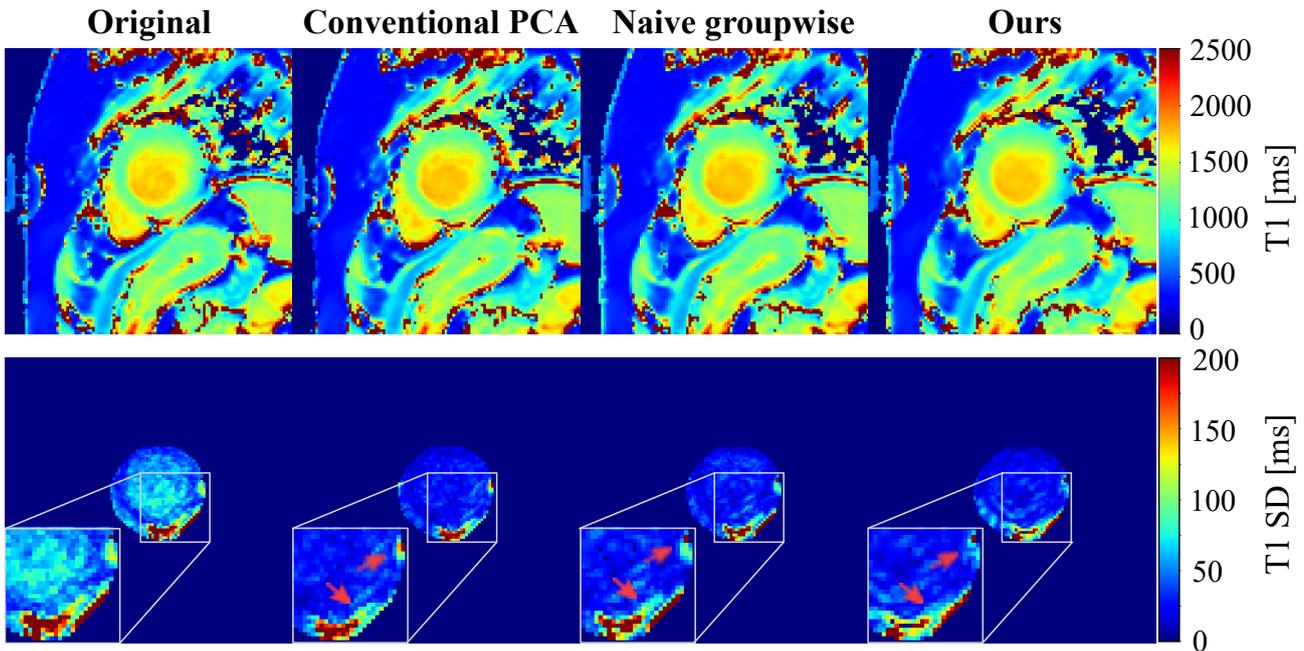


Figure 5. T1 fitting maps and SD errors for the MOLLI series. The first row shows the T1 maps, and the second row is the SD error. From left to right: original images, conventional PCA, naive groupwise, and our rPCA-based method. As the red arrows show, our method effectively corrects the motion in the boundary.

Table 2. Experiment results on T1 mapping of numerical phantoms. The RMSE measures the difference between the estimated T1 maps and real T1 maps, which means the lower the better. Our proposed method (w/ rPCA) outperforms the GroupRegnet\* in all metrics.

Method	RMSE ↓	SD diff ↑	$\mathcal{D}_{PCA}$ ↑
ConventionalPCA	48.42	<b>0.53</b>	<b>0.05</b>
GroupRegNet*	46.72	0.01	0.01
Ours (w/ rPCA)	<b>48.34</b>	0.14	0.02

we were constrained to generating only one basic  $T1$  map. In this case, we adopted the inversion time  $TI$  and learned deformation field of 11 real testing data and set the fitting parameter  $A(x, y) = 2, B(x, y) = 75 \times A(x, y)$  for each pixel to generate the synthetic data. The trained model based on the pre-contrast data was used to register the synthetic phantoms without any fine-tuning. By utilizing real groundtruth data, our proposed method demonstrated superior performance compared to the GroupRegNet\*, as evidenced by the results presented in Table 2. The  $SD$  diff results were consistent with the results on real data, ConventionalPCA gave best results because it was a subject-specific method. However, comparing the  $RMSE$  of ROI demonstrated that our proposed method performed similarly and even slightly better than *ConventionalPCA*.

#### 4. Conclusion and Discussion

In conclusion, we proposed a novel framework based on rPCA for robust motion correction of quantitative cardiac MRI. We aim for robust performance in the face of agnostic image contrast, which is typical of quantitative MRI. We showed that the introduction of rPCA, which separates low-rank and sparse components of baseline images, led to improved registration performance and facilitated the generalization of the trained network on out-of-domain data.

This work also compared the two commonly used metrics for groupwise registration, namely, NCC and NMI, and showed that NCC might give rise to potential artifacts in heart anatomy and quantitative mapping, both of which are detrimental to clinical application. Future investigations are warranted to focus not only on the performance of image registration but also on the fidelity of quantitative mapping.

Significantly, we validated our method using a realistic numerical phantom, which further proved the efficacy of our proposed approach. Noticeably, our method is additionally validated with a realistic numerical phantom and the results further prove the effectiveness of our proposed methods. While the synthetic data generated with MRXCAT2.0 already possesses a higher degree of realism compared to solely relying on physical anatomical data, in the future the incorporation of generative models, such as GANs, VAEs, and diffusion models, can yield even more diverse and realistic data. By utilizing these realistic and diverse digi-

tal phantoms, we not only establish a standardized dataset but also alleviate the challenges associated with acquiring labeled cardiac MRI images. This approach enables more comprehensive evaluations and facilitates the development of robust and generalizable registration techniques.

## References

- [1] Sahar Ahmad, Jingfan Fan, Pei Dong, Xiaohuan Cao, Pew-Thian Yap, and Dinggang Shen. Deep learning deformation initialization for rapid groupwise registration of inhomogeneous image populations. *Frontiers in Neuroinformatics*, 13:34, 2019. [2](#)
- [2] John Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007. [1](#)
- [3] Brian B Avants, Charles L Epstein, Murray Grossman, and James C Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, 2008. [3](#)
- [4] Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8):1788–1800, 2019. [4](#)
- [5] Mikael Brudfors, Yaël Balbastre, and John Ashburner. Groupwise multimodal image registration using joint total variation. In *Medical Image Understanding and Analysis: 24th Annual Conference, MIUA 2020, Oxford, UK, July 15-17, 2020, Proceedings 24*, pages 184–194. Springer, 2020. [1](#)
- [6] Stefano Buoso, Thomas Joyce, Nico Schulthess, and Sebastian Kozzerke. Mrxcat2. 0: Synthesis of realistic numerical phantoms by combining left-ventricular shape learning, biophysical simulations and tissue texture generation. *Journal of Cardiovascular Magnetic Resonance*, 25(1):25, 2023. [2](#), [5](#)
- [7] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):1–37, 2011. [2](#), [3](#)
- [8] Tongtong Che, Yuanjie Zheng, Jinyu Cong, Yanyun Jiang, Yi Niu, Wanzhen Jiao, Bojun Zhao, and Yanhui Ding. Deep group-wise registration for multi-spectral images from fundus images. *IEEE Access*, 7:27650–27661, 2019. [2](#)
- [9] Xiang Chen, Andres Diaz-Pinto, Nishant Ravikumar, and Alejandro F Frangi. Deep learning in medical image registration. *Progress in Biomedical Engineering*, 3(1):012003, 2021. [1](#)
- [10] Kelvin Chow, Jacqueline A Flewitt, Jordin D Green, Joseph J Pagano, Matthias G Friedrich, and Richard B Thompson. Saturation recovery single-shot acquisition (sasha) for myocardial t1 mapping. *Magnetic resonance in medicine*, 71(6):2082–2095, 2014. [2](#)
- [11] Albert de Roos and Charles B Higgins. Cardiac radiology: centenary review. *Radiology*, 273(2S):S142–S159, 2014. [1](#)
- [12] Bob D de Vos, Bas HM van der Velden, Jörg Sander, Kenneth GA Gilhuijs, Marius Staring, and Ivana Išgum. Mutual information for unsupervised deep learning image registration. In *Medical Imaging 2020: Image Processing*, volume 11313, pages 155–161. SPIE, 2020. [1](#), [3](#), [6](#)
- [13] Tobias Fechter and Dimos Baltas. One-shot learning for deformable medical image registration and periodic motion tracking. *IEEE transactions on medical imaging*, 39(7):2506–2517, 2020. [2](#)
- [14] Jiashi Feng, Huan Xu, and Shuicheng Yan. Online robust pca via stochastic optimization. *Advances in neural information processing systems*, 26, 2013. [3](#)
- [15] Qianjin Feng, Yujia Zhou, Xueli Li, Yingjie Mei, Zhentai Lu, Yu Zhang, Yanqiu Feng, Yaqin Liu, Wei Yang, and Wufan Chen. Liver dce-mri registration in manifold space based on robust principal component analysis. *Scientific reports*, 6(1):34461, 2016. [1](#), [2](#)
- [16] Xiujuan Geng, Gary E Christensen, Hong Gu, Thomas J Ross, and Yihong Yang. Implicit reference-based group-wise image registration and its application to structural and functional mri. *Neuroimage*, 47(4):1341–1351, 2009. [1](#)
- [17] Ricardo A Gonzales, Qiang Zhang, Bartłomiej W Papież, Konrad Werys, Elena Lukaschuk, Iulia A Popescu, Matthew K Burrage, Mayooraan Shanmuganathan, Vanessa M Ferreira, and Stefan K Piechnik. Moconet: robust motion correction of cardiovascular magnetic resonance t1 mapping using convolutional neural networks. *Frontiers in Cardiovascular Medicine*, page 1689, 2021. [2](#)
- [18] Jean-Marie Guyader, Wyke Huizinga, Dirk HJ Poot, Matthijs van Kranenburg, André Uitterdijk, Wiro J Niessen, and Stefan Klein. Groupwise image registration based on a total correlation dissimilarity measure for quantitative mri and dynamic imaging data. *Scientific reports*, 8(1):13112, 2018. [2](#)
- [19] Valentin Hamy, Nikolaos Dikaios, Shonit Punwani, Andrew Melbourne, Arash Latifoltojar, Jessica Makanyanga, Manil Chouhan, Emma Helbren, Alex Menys, Stuart Taylor, et al. Respiratory motion correction in dynamic mri using robust data decomposition registration—application to dce-mri. *Medical image analysis*, 18(2):301–313, 2014. [2](#)

- [20] Wyke Huizinga, Dirk HJ Poot, J-M Guyader, Remy Klaassen, Bram F Coolen, Matthijs van Kranenburg, RJM Van Geuns, André Uitterdijk, Mathias Polfliet, Jef Vandemeulebroucke, et al. Pca-based groupwise image registration for quantitative mri. *Medical image analysis*, 29:65–78, 2016. 1, 2, 4
- [21] Peter Kellman, Andrew E Arai, and Hui Xue. T1 and extracellular volume mapping in the heart: estimation of error maps and the influence of noise on precision. *Journal of Cardiovascular Magnetic Resonance*, 15(1):1–12, 2013. 1, 4
- [22] Arno Klein, Jesper Andersson, Babak A Ardekani, John Ashburner, Brian Avants, Ming-Chang Chiang, Gary E Christensen, D Louis Collins, James Gee, Pierre Hellier, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *Neuroimage*, 46(3):786–802, 2009. 1
- [23] Andrew C Larson, Richard D White, Gerhard Laub, Elliot R McVeigh, Debiao Li, and Orlando P Simonetti. Self-gated cardiac cine mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 51(1):93–102, 2004. 2
- [24] Yuze Li, Chunyan Wu, Haikun Qi, Dongyue Si, Haiyan Ding, and Huijun Chen. Motion correction for native myocardial t1 mapping using self-supervised deep learning registration with contrast separation. *NMR in Biomedicine*, 35(10):e4775, 2022. 2
- [25] Timo Makela, Patrick Clarysse, Outi Sipila, Nicoleta Pauna, Quoc Cuong Pham, Toivo Katila, and Isabelle E Magnin. A review of cardiac image registration methods. *IEEE Transactions on medical imaging*, 21(9):1011–1021, 2002. 1
- [26] Daniel R Messroghli, Aleksandra Radjenovic, Sebastian Kozerke, David M Higgins, Mohan U Sivanathan, and John P Ridgway. Modified look-locker inversion recovery (molli) for high-resolution t1 mapping of the heart. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 52(1):141–146, 2004. 1, 2, 4
- [27] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015. 4
- [28] Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8):712–721, 1999. 3, 4
- [29] W Paul Segars, G Sturgeon, S Mendonca, Jason Grimes, and Benjamin MW Tsui. 4d xcat phantom for multimodality imaging research. *Medical physics*, 37(9):4902–4915, 2010. 2
- [30] Qian Tao, Pieter van der Tol, Floris F Berendsen, Elisabeth HM Paiman, Hildo J Lamb, and Rob J van der Geest. Robust motion correction for myocardial t1 and extracellular volume mapping by principle component analysis-based groupwise image registration. *Journal of Magnetic Resonance Imaging*, 47(5):1397–1405, 2018. 1, 2
- [31] Norbert M Wilke, Michael Jerosch-Herold, Andrey Zenovich, and Arthur E Stillman. Magnetic resonance first-pass myocardial perfusion imaging: clinical validation and future applications. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 10(5):676–685, 1999. 2
- [32] Hui Xue, Saurabh Shah, Andreas Greiser, Christoph Guetter, Arne Littmann, Marie-Pierre Jolly, Andrew E Arai, Sven Zuehlsdorff, Jens Guehring, and Peter Kellman. Motion correction for myocardial t1 mapping using image registration with synthetic image estimation. *Magnetic resonance in medicine*, 67(6):1644–1655, 2012. 1
- [33] Yunlu Zhang, Xue Wu, H Michael Gach, Harold Li, and Deshan Yang. Groupregnet: a groupwise one-shot deep learning-based 4d image registration method. *Physics in Medicine & Biology*, 66(4):045030, 2021. 2, 4, 5
- [34] Tianyi Zhou and Dacheng Tao. Godec: Randomized low-rank & sparse matrix decomposition in noisy case. In *Proceedings of the 28th International Conference on Machine Learning, ICML 2011*, 2011. 6

# 3

## Quantitative Cardiac MRI

Magnetic Resonance Imaging (MRI) employs a powerful magnet and radio waves to examine the human body and generate visual representations. Unlike various other imaging techniques, MRI does not employ ionizing radiation, which is associated with potential long-term risks. By producing high-resolution images, MRI proves particularly advantageous in intricate scenarios, notably for cardiac evaluations such as the detection of coronary heart disease and cardiac tumors. Furthermore, MRI offers a commendable safety profile, making it especially valuable for patients requiring multiple scans over an extended period, including individuals with complex congenital heart conditions. Cardiac MRI means the images relating to the heart.

In this chapter, we present the background of quantitative cardiac MRI, the fundamental theory of T1 mapping and the MOLLI sequences used in our project.

### 3.1. Quantitative Cardiac MRI

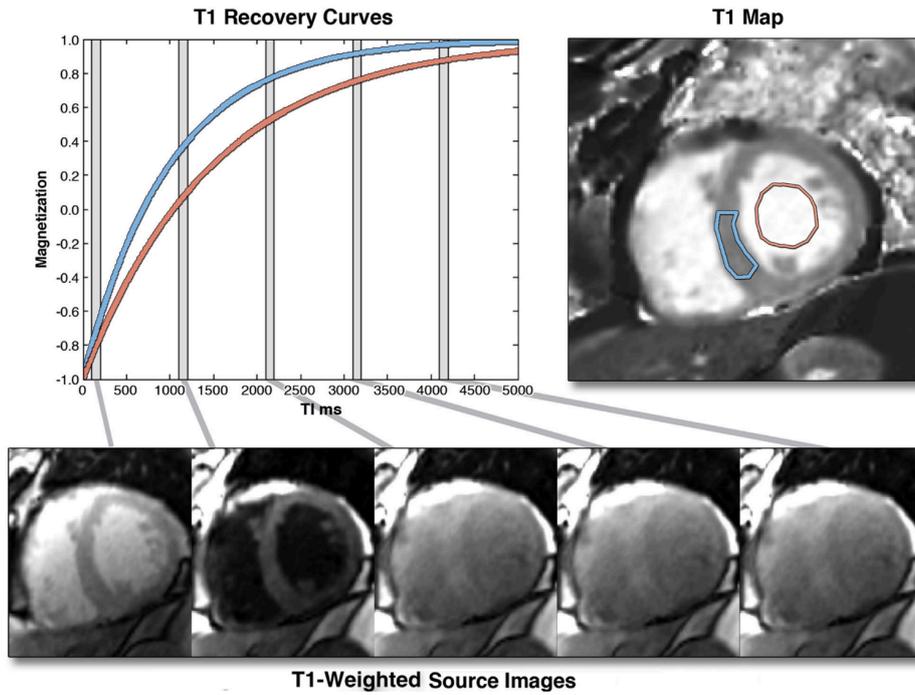
Conventional Magnetic Resonance Imaging (MRI) primarily relies on the acquisition of weighted images, which primarily aim to enhance local image contrast. However, these conventional MRI techniques often struggle to detect significant morphological abnormalities or focal irregularities that result in regional variations in signal intensities [14]. This limitation arises due to their insensitivity towards subtle global changes. Additionally, Carlo's report [14] highlights that conventional MRI techniques face challenges in distinguishing between different physiological and pathological patterns if their alterations in image contrast are similar.

To address these limitations, quantitative MRI techniques are proposed, aiming to provide information about the local microstructural environment of protons. Quantitative MRI techniques offer specific physical parameters related to the nuclear spin of protons in water, such as relaxation times [20]. A significant advantage of quantitative MRI is its ability to establish normative measurements within a healthy population. This allows for enhanced sensitivity in clinical MRI by comparing measurements from a single subject to these novel reference values. By quantifying these parameters, quantitative MRI facilitates a more comprehensive understanding of the underlying tissue characteristics. This approach proves particularly valuable in monitoring subtle changes attributed to disease progression or remission.

For diagnosis, quantitative MRI often refers to true relaxation times T1, T2 or T2\* [9]. In this project, according to our dataset, we use the estimated T1 mapping for evaluation. Therefore, we will mainly introduce the technical details regarding T1 mapping.

### 3.1.1. T1 mapping

Cardiac magnetic resonance imaging is sensitive to the soft tissue image contrast, providing insights into physiology and pathophysiology. The T1 relaxation time serves as a crucial biomarker for various pathological conditions, measuring how fast the nuclear spin magnetization returns to its equilibrium state after a radiofrequency (RF) pulse in the MRI scanner. The resulting T1 map is a color-encoded image, where each pixel value represent the T1 relaxation time for each voxel. High-quality T1 maps can reveal small variations of T1 within the heart to highlight tissue pathology.



**Figure 3.1:** Magnetization Inversion Recovery for  $T_1$  mapping [19]. The T1-Weighted source images (**bottom-row**) are acquired using different  $T_1$  weightings during the scan, taken at different inversion time (TI) after an inversion pulse. The  $T_1$  recovery curves (**top-left**) shows 2 curves for two example regions, orange curve for the blood pool region on  $T_1$  map (**top-right**) and blue curve for septal region.

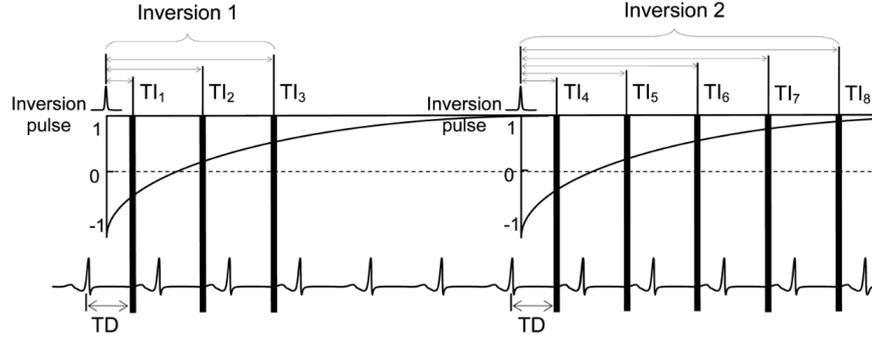
Accurate measurement and mapping of  $T_1$  relaxation time require a series of images with different T1 weightings to fit the signal intensities of the images to the equation for  $T_1$  relaxation. As shown in Figure 3.1, a sequence of  $T_1$ -weighted source images are acquired at different times after the pulse. For the pixel at every position, the signal intensity can be fit to the equation:

$$y(\theta) = A - Be^{-t/T_1}, \quad (3.1)$$

where  $\theta = (A, B, T_1)$ ,  $A, B$  are fitting parameters and  $t$  represents the time after the preparation.

### 3.1.2. Modified Look-Locker Pulse Sequence

The modified Look-Locker sequence (MOLLI) [13] is a widely used and advanced clinical technique for  $T_1$  mapping, which is also the sequences we used in our experiments. In MOLLI, multiple inversions with slightly different  $TIs$  are used to evenly sample the recovery curve and the data is acquired at end-diastole when the heart is reasonably stationary. The Figure 3.2



**Figure 3.2:** MOLLI sequence scheme shows two sets of inversions were performed with increasing inversion time (TI) [23]. The trigger delay (TD) is used to acquire the image at end-diastole phase.

shows an example of MOLLI sequences, where additional inversions are employed, and all images are acquired at the same end-diastolic phase.

Compared to the  $T_1$  recovery curves shown in Figure 3.1, MOLLI technique affect the  $T_1$  recovery curve, resulting in an apparent  $T_1$ , denoted as  $T_1^*$ . The estimation of  $T_1$  relaxation will be:

$$y(\theta) = A - Be^{-t/T_1^*}, \quad (3.2)$$

$$T1 = \left(\frac{B}{A} - 1\right) \times T_1^*, \quad (3.3)$$

where  $\theta = (A, B, T_1^*)$ ,  $A, B$  are still the fitting parameters and  $t$  denotes the inversion time.

The MOLLI technique reduces the influence of cardiac motion as all images are acquired during the same cardiac phase. However, unintended patient movements and breathing introduce errors in pixel-wise  $T_1$  estimation, resulting in inaccurate final  $T_1$  maps. Therefore, motion correction is essential to generate reliable  $T_1$  mapping.

# 4

## Deformable Image Registration

Deformable image registration (DIR) is a popular topic in medical imaging analysis, aiming to estimate the non-linear transformation between a pair or a series of images to align them. The term deformable instead of linear or global is used because the observed motions are non-linear dense transformation. Let  $\Omega$  be an  $n$ -D spatial domain and  $\mathbf{x}$  be the corresponding  $n$ -D coordinate,  $I^N = \{I_i(x) \in \mathbb{R}^\Omega | i = 1, \dots, N\}$  denote a series of  $N$  images. When a spatial coordinate  $\mathbf{x}$  does not correspond to the same anatomical location in each of the images, it becomes necessary to correct for motion or geometric distortion using DIR methods. A set of deformation fields  $T(\mathbf{x})^N = \{T_i(\mathbf{x}) \in \mathbb{R} | i = 1, \dots, N\}$  is learned and the motion-corrected images are generated using  $T_i(x) \circ I_i(x)$ , where  $\circ$  represents the spatial transformation. The registration framework can be categorized into pairwise and groupwise methods [11].

This chapter introduces main categories of deformable image registration frameworks, widely used similarity metric and the theory of robust principal component analysis used in our project.

### 4.1. DIR frameworks: Pairwise and groupwise

For pairwise methods, the transformation is modeled by a set of transform parameters  $\boldsymbol{\mu}$ . Each image  $I_i(x)$  has a corresponding deformation field  $T_i(\mathbf{x}; \boldsymbol{\mu}_i)$ . One reference image  $I_R(\mathbf{x})$  is selected and all other images  $I_i(\mathbf{x})$ ,  $i \neq R$ , are registered to the reference  $I_R(\mathbf{x})$ . The pairwise DIR aims to minimize a dissimilarity metric  $\mathcal{D}$  in terms of  $\boldsymbol{\mu}_i$  [11]:

$$\hat{\boldsymbol{\mu}}_i = \arg \min_{\boldsymbol{\mu}_i} \mathcal{D}(\boldsymbol{\mu}_i; I_R(\mathbf{x}), I_i(\mathbf{x})), \quad (4.1)$$

which is calculated for all images  $i \neq R$ . The dissimilarity metric  $\mathcal{D}$  can be relevant definition measuring the dissimilarity of reference image  $I_R(\mathbf{x})$  and registered images  $T_i(\mathbf{x}, \boldsymbol{\mu}_i) \circ I_i(\mathbf{x})$ .

For groupwise methods, the series of images  $I_i(\mathbf{x})$  are registered simultaneously to a common space. The deformation field is modeled by transform parameter  $\boldsymbol{\mu}$ , where  $\boldsymbol{\mu}$  contains all  $\boldsymbol{\mu}_i$ . The groupwise registration can be formulated as the minimization of a dissimilarity metric  $\mathcal{D}$  in terms of  $\boldsymbol{\mu}$ :

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}_i} \mathcal{D}(\boldsymbol{\mu}; I(\mathbf{x})), \quad (4.2)$$

where the parameters  $\boldsymbol{\mu}_i$  are simultaneously optimized for all  $i$  [24].

## 4.2. DIR frameworks: Conventional and Deep Learning-based

Most conventional DIR algorithms optimize a transformation by iteratively minimizing an energy function [18]. The energy function typically consists of two components:

$$\mathcal{L}_{sim}(I, T) + \mathcal{L}_{smooth}(T), \quad (4.3)$$

where  $\mathcal{L}_{sim}(I, T)$  measures the level of alignment and  $\mathcal{L}_{smooth}(T)$  is the regularization term, ensuring spatial smoothness and homogeneity of the deformation field  $T$ . Here the deformation field  $T$  is unique for each subject as the energy function is optimized subject-specific. Each image pair has its own unique deformation field, making the optimization process subject-specific. However, this approach becomes time-consuming, particularly for high-resolution images.

In recent years, deep learning-based DIR frameworks have emerged and give promising results, offering an alternative to traditional methods. These frameworks replace the subject-specific parameters with shared parameters which are optimized by a global optimization. The neural network is trained on a dataset of image pairs, enabling the fast generation of deformation fields for a given image during inference. In deep learning-based method, we model a function  $g_{\theta}(I_f, I_m) = \mu$ , where  $\theta$  represents the network parameters and output  $\mu$  estimates the deformation field's parameters. The network takes a pair of images  $I_f$  (fixed image) and  $I_m$  (moving image) as the input (for pairwise methods as an example, while groupwise image registration only requires a series of input image). The deformation field  $T$  is calculated as follows :

$$T = Id + \mu, \quad (4.4)$$

where  $Id$  is the identity transform. Once the deformation field is obtained, the moving image  $I_m$  is warped to  $T \circ I_m$  using a spatial transformation function. To find optimal parameters  $\hat{\theta}$  unsupervised loss function is defined as follows [2]:

$$\mathcal{L}_{sim}(I_f, T \circ I_m) + \lambda \mathcal{L}_{smooth}(T), \quad (4.5)$$

where  $I_f$  and  $I_m$  are two continuous frames from an image series, and in each step, two frames are randomly selected from the dataset.

Deformable image registration, a critical task in medical image analysis, employs various neural network architectures for optimizing the deformation parameters. A diverse range of architectures have demonstrated efficacy in this field. For instance, the popular UNet architecture [15] including encoder and decoder sections with skip connections, widely used in medical image analysis, has been successfully applied in VoxelMorph [2]. Alternatively, transformer-based architectures, such as the Swin-Voxelmorph [25], offer another viable option. These transformer-based models, inspired by the success of the transformer architecture [12] in various domains, have shown promise in deformable image registration tasks. The UNet architecture implemented in our work is shown in Figure 4.1. The encoder in the left section uses several convolutions with a stride of 2 to capture the hierarchical features and reduce the spatial dimension. The decoder section leverages the upsampling and skip connection to propagate the features.

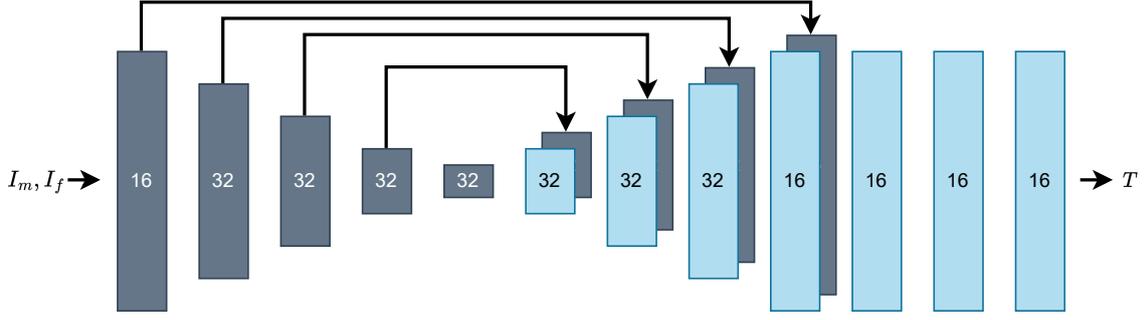


Figure 4.1: UNet architecture [2].

## 4.3. Similarity Criterion

The similarity term penalizes the differences and several popular functions exist. In the following section, mean square error, cross correlation, mutual information and total correlation are introduced.

### 4.3.1. Mean Square Error

Mean Square Error (MSE) measures the voxel-wise/pixel-wise mean squared difference and is useful when  $I_f$  and  $I_m$  have similar intensity distribution and local contrast [8]:

$$MSE(I_f, T \circ I_m) = \frac{1}{\Omega} \sum_{x \in \Omega} \|I_f(x) - [T \circ I_m](x)\|^2. \quad (4.6)$$

### 4.3.2. Cross Correlation

Cross Correlation (CC) is a similarity metric more robust to intensity variations [1], only estimating the local image average variance. We define  $\hat{I}_f(x)$  and  $[T \circ \hat{I}_m](x)$  as the local mean intensity image, which is computed over a local  $n^D$  window centered at each position  $x$ . We use a variable to represent the image with its local mean subtracted as  $\bar{I}(x) = I_f(x) - \hat{I}_f(x)$  and  $\bar{J}(x) = [T \circ I_m](x) - [T \circ \hat{I}_m](x)$  for simplicity. The cross-correlation is defined as follow: [8]:

$$CC(\bar{I}(x), \bar{J}(x)) = \frac{\langle \bar{I}, \bar{J} \rangle^2}{\langle \bar{I} \rangle \langle \bar{J} \rangle}, \quad (4.7)$$

where  $\langle \rangle$  means the inner product computing over a  $n^D$  window.

### 4.3.3. Mutual Information

Mutual Information(MI) is a voxel-based similarity measurement which is insensitive to the intensity changes. Derived from information theory, MI quantifies the statistical dependency between two images, making it a valuable tool for alignment purposes. The MI between two images,  $I_m$  and  $I_f$ , is computed as follows [16]:

$$MI(I_m, I_f) = H(I_m) + H(I_f) - H(I_m, I_f), \quad (4.8)$$

where  $H(I_m), H(I_f)$  denotes the marginal entropy of image  $I_m$  and  $I_f$ , respectively. The term  $H(I_m, I_f)$  denotes their joint entropy, which is calculated from the joint histogram of  $I_m$  and

$I_f$ . The joint entropy is calculated from the joint histogram of image  $I_m$  and  $I_f$ . The entropy of an image  $A$  is defined as:

$$H(A) = - \sum_a p_A(a) \log p_A(a), \quad (4.9)$$

and the joint entropy between two images  $A$  and  $B$  is:

$$H(A, B) = - \sum_{a,b} p_{A,B}(a, b) \log p_{A,B}(a, b). \quad (4.10)$$

MI measures the amount of shared information between the two images. Consequently, when two images are well aligned, the MI value increases, indicating a higher degree of mutual information between them.

#### 4.3.4. Total Correlation

Total Correlation(TC) is used in conventional DLR framework as groupwise dissimilarity metric. The image series  $I_i$  can be represented as a  $G \times N$  matrix  $M$ , where  $N$  is the number of frames in the series,  $G$  is the number of voxels in one image  $I_i$ . Each row of  $M$  denotes a data point in a  $N$ -dimensional space.

The correlation matrix of the data points in  $M$  is defined as:

$$K = \frac{1}{N-1} \Sigma^{-1} (M - \bar{M})^T (M - \bar{M}) \Sigma^{-1}, \quad (4.11)$$

where  $\bar{M}$  is a average matrix that each of its column is the column-wise average of  $M$  and  $\Sigma$  is a diagonal matrix that its diagonal elements are the standard deviation of the column of  $M$ . The diagonal matrix  $\Sigma$  is used to make this metric robust to the linear intensity scalings and offsets. The idea is that once the images are aligned, the higher eigenvalues increase.

The dissimilarity metric based on total correlation  $\mathcal{D}_{TC}$  measuring the amount of information shared between any subset of image  $I^N$ , is defined as follows [10]:

$$\begin{aligned} \mathcal{D}_{TC}(I) &= H(I) - \sum_{i=1}^N H(I_i) \\ &= \frac{N}{2} + \frac{N}{2} \ln(2\pi) + \frac{1}{2} \ln(\det(\mathbf{K})) - \sum_{i=1}^N H(I_i) \\ &= \frac{N}{2} + \frac{N}{2} \ln(2\pi) + \frac{1}{2} \ln(\det(\mathbf{K})) - \sum_{i=1}^N \left( \frac{1}{2} + \frac{1}{2} \ln(2\pi) \right) \\ &= \frac{1}{2} \ln(\det(\mathbf{K})) \\ &= \frac{1}{2} \sum_{i=1}^N \ln \lambda_i, \end{aligned} \quad (4.12)$$

where  $\det(\mathbf{K}) = \prod_{i=1}^N \lambda_j$  and  $\lambda_j$  represents the  $j^{\text{th}}$  eigenvalue of correlation matrix  $\mathbf{K}$ .

## 4.4. B-spline Transform Model

The linear transformation is the simplest transformation model, describing the transform direction and size. For example, project the point  $(x_1, y_1, z_1)$  to  $(x_1 + \mu_x, y_1 + \mu_y, z_1 + \mu_z)$  in 3-D spatial domain. The linear model can only capture the global image motion, where the free form deformation (FFD) model based on B-spline provides a local deformation model that offers more flexibility and precision.

The B-spline model treats local feature separately because it generates spline basis functions for each feature independently. Consider a 3D FFD B-spline transform as an example, where the spatial domain  $\Omega = \{(x, y, z) | 0 \leq x \leq X, 0 \leq y \leq Y, 0 \leq z \leq Z\}$  and we use  $\Psi_{i,j,k}$  to denote a control point. The FFD transformation is expressed using a cubic B-spline as follows[22]:

$$T_{nonrigid}(x, y, z) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(u) B_m(v) B_n(w) \Psi_{i+l, j+m, k+n}. \quad (4.13)$$

Here, the B-spline model defines the corresponding kernel functions as:

$$\begin{aligned} B_0(u) &= (1 - u)^3/6, \\ B_1(u) &= (3u^3 - 6u^2 + 4)/6, \\ B_2(u) &= (-3u^3 + 3u^2 + 3u + 1)/6, \\ B_3(u) &= u^3/6. \end{aligned} \quad (4.14)$$

By parameterizing the B-spline transformation with control points, the FFD model allows for local modifications in number of control point. This approach proves to be effective and naturally smooths the deformation field, enabling accurate modeling of local deformations.

In contrast to linear transformation models that primarily capture global image motion, the FFD model based on B-spline transforms offers enhanced flexibility and the ability to represent intricate local deformations. Consequently, this model is particularly suitable for medical image registration tasks where precise alignment of local anatomical structures is critical.

## 4.5. Robust Principal Component Analysis

The fact that high-dimensional data usually lies on some low-dimensional manifold [3] makes it possible to decompose the large matrix  $M$  into two sub-matrix:

$$M = L_0 + S_0, \quad (4.15)$$

where the  $L_0$  is a low-rank matrix and  $S_0$  is sparse. Compared the classical *Principal Component Analysis* (PCA) which seeks the best (in an  $\ell^2$  sense) rank- $k$  estimate of  $L_0$  by solving [7]:

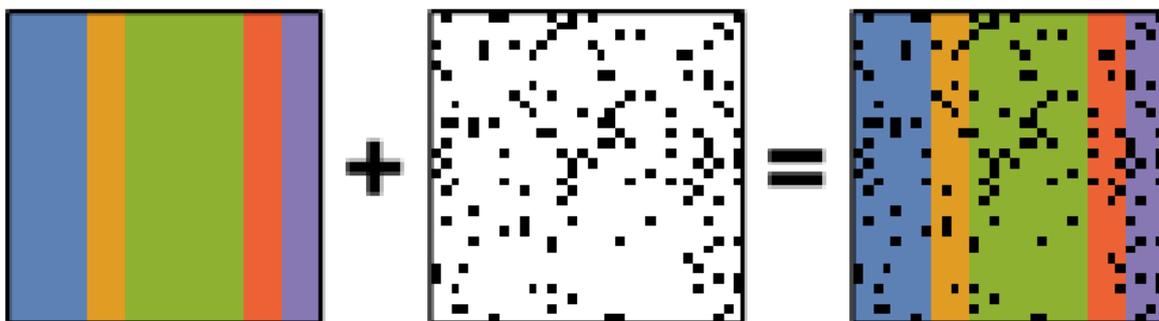
$$\begin{aligned} &\text{minimize} \quad \|M - L\| \\ &\text{subject to} \quad \text{rank}(L) \leq k. \end{aligned} \quad (4.16)$$

This problem is effectively solved through the singular value decomposition (SVD), decomposing the matrix  $M = L_0 + N_0$ , where  $N_0$  is a small perturbation matrix. In *Robust Principal Component Analysis* (rPCA), instead of using the small noise term  $N_0$ , a sparse matrix  $S_0$  is estimated. The rPCA is solved using Principal Component Pursuit (PCP) estimate solving

$$\begin{aligned} &\text{minimize} \quad \|L\|_* + \lambda \|S\|_1 \\ &\text{subject to} \quad L + S = M. \end{aligned} \quad (4.17)$$

The  $\|A\|_* := \sum_i \sigma_i(A)$  denotes the nuclear norm of the matrix  $A$ , i.e. the sum of the singular values of  $A$ ,  $\|A\|_1 = \sum_{ij} |A_{ij}|$  denotes the  $\ell_1$ -norm of  $A$  seen as a long vector in  $\mathbb{R}^{n_1 \times n_2}$  [5].

PCA is known to be sensitive to outliers in data while rPCA is more robust with robust errors. By robustly decomposing data into low-rank and sparse components, as shown in Figure 4.2, rPCA enables the extraction of meaningful information from noisy or corrupted datasets, enhancing the reliability and interpretability of the results.



**Figure 4.2:** Schematic diagram of the robust PCA problem, which combines a low-rank matrix with sparse errors. Robust PCA aims to decompose the matrix back into these two components. [6]

# 5

## Numerical Phantoms Synthesis

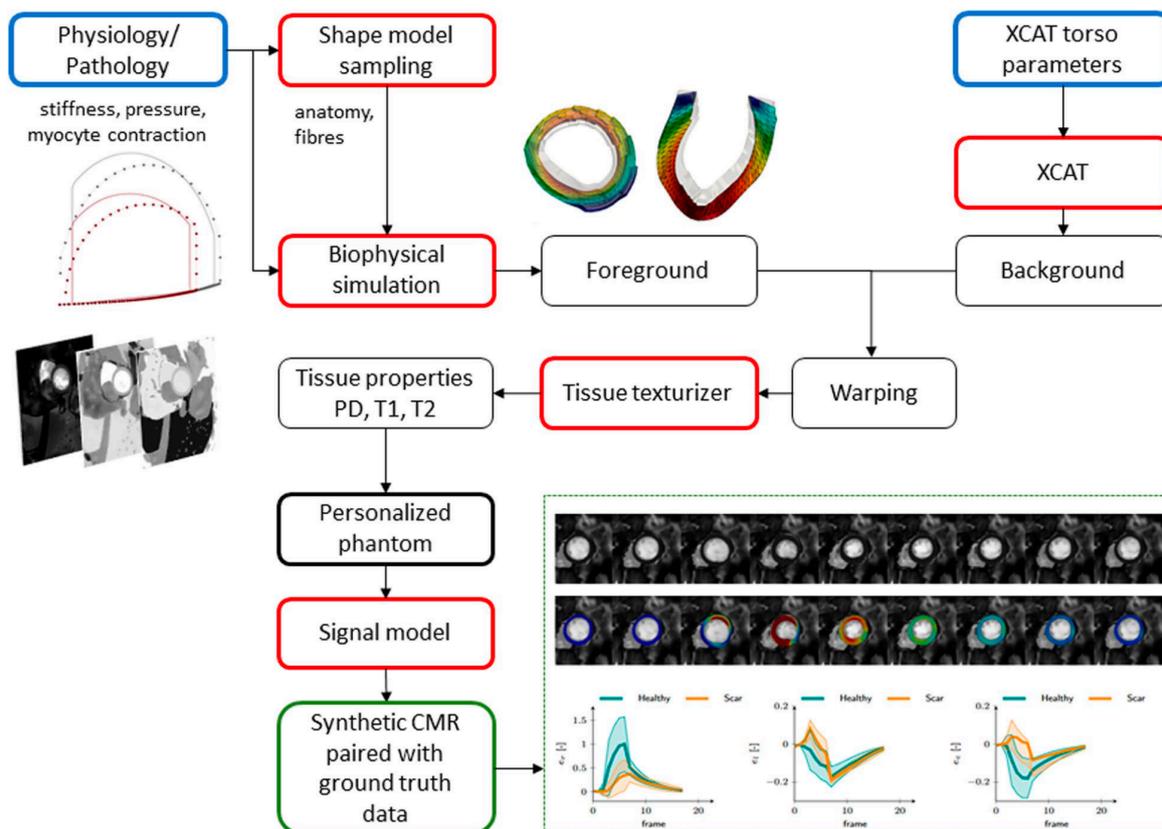
MRI image acquisition, particularly for cardiac magnetic resonance (CMR) imaging, is known to be expensive and time-consuming. To mitigate these limitations, the utilization of in-silico phantoms has emerged as a powerful tool for testing and validating various tasks in MRI, such as image reconstruction, image registration, and image segmentation. Generating synthetic images from digital phantoms offers several advantages, including the availability of corresponding anatomical labels and functional ground truth, facilitating evaluation and analysis.

Voxel-based methods in the context of phantom synthesis rely on labeled voxelized representations derived from real patient data, yielding highly realistic phantoms. However, these methods often suffer from limited generalizability due to their dependence on specific patient datasets. On the other hand, analytical models are based on mathematical descriptions of tissue structures, which enable accurate phantoms but may sacrifice some realism. To overcome these inherent limitations, hybrid methods have been proposed, combining the strengths of voxel-based and analytical approaches.

In our research project, we have chosen to utilize the MRXCAT2.0 framework [4], an new version of the previous MRXCAT software [21], for digital phantom synthesis in the field of MRI. MRXCAT2.0 provides a comprehensive and versatile platform that allows for the generation of synthetic images that are both anatomically accurate and realistic, facilitating the evaluation of various imaging techniques.

MRXCAT2.0 is a hybrid method aiming to address the two limitations: lack of variability and realism. It combines a statistical model to generate realistic left ventricular anatomy and function with the XCAT model [17] for the surrounding tissue structure. Additionally, a trained neural network is utilized to generate realistic tissue maps and relaxation times, among other parameters. The overall pipeline from MRXCAT2.0 can be seen in Figure 5.1.

The input of the MRXCAT2.0 includes two parts: one is the selection of physiological characteristics of anatomy, another is the parameters for the XCAT phantom. The first part is employed to generate the foreground through the biophysical simulation and the second part is used to generate the background, defining the torso anatomy and the displacement field. The foreground and background components are subsequently combined and undergo warping and tissue texturization processes to calculate tissue properties. The ground truth data in our work, specifically the T1 map, is generated based on these tissue properties. While MRXCAT2.0 typically utilizes the balanced steady-state free precession (bSSFP) signal model, we have replaced it with the synthetic process described in our previous work Chapter 2. In our project,



**Figure 5.1:** MRXCAT2.0 workflow [4].

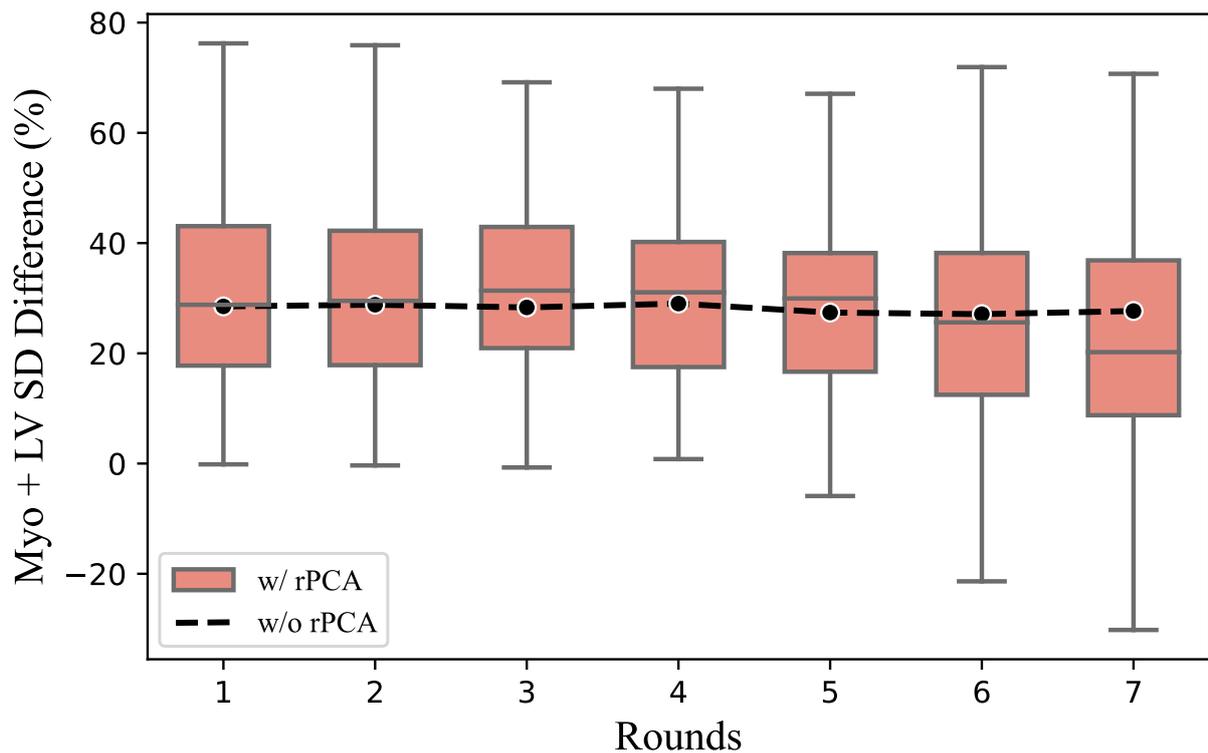
our personalized signal model (The  $T_1$  exponential fitting curve) to generate the synthetic  $T_1$  maps.

# 6

## Additional Experiments

In this chapter, we present the ablation studies that are not written in the academic paper in Chapter 2.

### 6.1. Round Selection



**Figure 6.1:** SD difference changes over rounds on **pre-contrast (in-domain)** data optimized using normalized mutual information. The SD difference compares the decrease of SD error in each round to the original data, where a larger difference indicates better registration. The performance of our method (w/ rPCA) improves at the beginning and decreases after a few rounds, which means it successfully corrects most of the motion through multi-rounds. The naive method (w/o rPCA) stays stable over rounds and performs worse than our method in our selected round (round 3).

## 6.2. Comparison of different similarity metrics

### 6.2.1. SD Difference

Different similarity metrics (NMI, NCC and TC) are compared according to the SD difference on our proposed rPCA-based method. The results are shown in Table 6.1 on pre-contrast data and Table 6.2 on post-contrast data. The results on both in-domain and out-of-domain data indicate our proposed method perform best using normalized mutual information.

**Table 6.1:** SD difference changes over rounds on **pre-contrast (in-domain)** data optimized using on normalized mutual information (NMI), normalized cross correlation (NCC) and total correlation (TC).

Round	1	2	3	4	5	6	7
NMI	0.3262	0.3206	<b>0.3401</b>	0.319	0.2973	0.2643	0.2280
NCC	0.3245	0.3214	0.2978	0.2758	0.2589	0.2471	0.2437
TC	0.3312	0.3353	0.331	0.3165	0.2998	0.2781	0.2531

**Table 6.2:** SD difference changes over rounds on **post-contrast (out-of-domain)** data optimized using on normalized mutual information (NMI), normalized cross correlation (NCC) and total correlation (TC).

Round	1	2	3	4	5	6	7
NMI	0.1045	0.1325	<b>0.1756</b>	0.1615	0.1779	0.1638	0.1658
NCC	0.0995	0.1196	0.1146	0.1154	0.1181	0.1009	0.0978
TC	0.1123	0.1344	0.1593	0.1895	0.1880	0.1828	0.1674

### 6.2.2. Examination of potential artefacts

To evaluate whether undesirable deformation within the ROI exists, in addition to the correlation plot shows in Chapter 2, we measure the average T1 value of estimated T1 map within the ROI. The large shift from the original average value might indicate the unintentional motion artefacts. The results of each subject and the average value are shown in Table 6.3, where the normalized mutual information contains the smallest shift from the original data. Combined with the results in Section 6.2.1, normalized mutual information is finally selected as the similarity metric in our project.

**Table 6.3:** Average T1 value of estimated T1 maps. Each row indicates averaged T1 value using different metrics. The orig indicates the T1 mapping without any registration, normalized mutual information (NMI), normalized cross correlation (NCC) and total correlation (TC) are listed. The last row is the averaged value over 20 test subjects. The average value using mutual information is the nearest to the original value.

Subject	Orig	NMI	NCC	TC
MOLLI_0379217_20131018_1	1486.58	1349.83	1341.26	1360.14
MOLLI_0387056_20140416_3	1357.46	1328.84	1324.90	1346.26
MOLLI_0387056_20140423_1	1452.84	1452.52	1503.66	1442.63
MOLLI_0440522_20140409_2	1261.96	1446.33	1438.62	1460.58
MOLLI_1328040_20131002_1	1467.56	1467.72	1491.55	1503.75
MOLLI_1328081_20140129_1	1221.37	1218.37	1138.46	1244.15
MOLLI_1351763_20131030_1	1461.45	1356.11	1369.45	1502.56
MOLLI_2265167_20140212_2	287.95	279.86	284.98	274.62
MOLLI_2550514_20140219_1	1318.23	1360.93	1366.55	1369.74
MOLLI_3355043_20140305_2	1448.49	1469.61	1456.02	1478.71
MOLLI_3386874_20140528_1	1524.44	1499.78	1474.46	1536.80
MOLLI_4143835_20140402_2	1496.25	1443.78	1449.47	1456.87
MOLLI_5179318_20131106_2	1472.56	1568.73	1582.90	1564.61
MOLLI_5639272_20131204_3	1664.31	1452.76	1445.87	1466.47
MOLLI_6157978_20140115_2	1495.37	1448.45	1445.78	1464.35
MOLLI_7324776_20140416_1	1194.77	1342.65	1308.72	1373.07
MOLLI_8101167_20140611_2	1346.84	1237.35	1229.46	1237.49
MOLLI_8559809_20140402_3	1385.99	1379.78	1377.53	1397.30
MOLLI_9034739_20140528_2	1326.32	1476.25	1472.23	1482.35
MOLLI_9353628_20140129_2	1450.30	1506.55	1523.96	1509.32
Average	<b>1356.05</b>	<b>1354.30</b>	1351.29	1373.58

# References

- [1] Brian B Avants et al. “Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain”. In: *Medical image analysis* 12.1 (2008), pp. 26–41.
- [2] Guha Balakrishnan et al. “Voxelmorph: a learning framework for deformable medical image registration”. In: *IEEE transactions on medical imaging* 38.8 (2019), pp. 1788–1800.
- [3] Mikhail Belkin and Partha Niyogi. “Laplacian eigenmaps for dimensionality reduction and data representation”. In: *Neural computation* 15.6 (2003), pp. 1373–1396.
- [4] Stefano Buoso et al. “MRXCAT2. 0: Synthesis of realistic numerical phantoms by combining left-ventricular shape learning, biophysical simulations and tissue texture generation”. In: *Journal of Cardiovascular Magnetic Resonance* 25.1 (2023), p. 25.
- [5] Emmanuel J Candès et al. “Robust principal component analysis?” In: *Journal of the ACM (JACM)* 58.3 (2011), pp. 1–37.
- [6] Francisco De La Peña et al. “hyperspy/hyperspy: Release v1. 7.3”. In: *Zenodo* ().
- [7] Carl Eckart and Gale Young. “The approximation of one matrix by another of lower rank”. In: *Psychometrika* 1.3 (1936), pp. 211–218.
- [8] Koen AJ Eppenhof et al. “Deformable image registration using convolutional neural networks”. In: *Medical Imaging 2018: Image Processing*. Vol. 10574. SPIE. 2018, pp. 192–197.
- [9] Daniel Gräfe et al. “Quantitative T1 mapping of the normal brain from early infancy to adulthood”. In: *Pediatric radiology* 51 (2021), pp. 450–456.
- [10] Jean-Marie Guyader et al. “Groupwise image registration based on a total correlation dissimilarity measure for quantitative MRI and dynamic imaging data”. In: *Scientific reports* 8.1 (2018), p. 13112.
- [11] Wyke Huizinga et al. “PCA-based groupwise image registration for quantitative MRI”. In: *Medical image analysis* 29 (2016), pp. 65–78.
- [12] Ze Liu et al. “Swin transformer: Hierarchical vision transformer using shifted windows”. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2021, pp. 10012–10022.
- [13] Daniel R Messroghli et al. “Modified Look-Locker inversion recovery (MOLLI) for high-resolution T1 mapping of the heart”. In: *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 52.1 (2004), pp. 141–146.
- [14] Carlo Pierpaoli. “Quantitative brain MRI”. In: *Topics in magnetic resonance imaging: TMRI* 21.2 (2010), p. 63.

- [15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer. 2015, pp. 234–241.
- [16] Daniel Rueckert et al. “Nonrigid registration using free-form deformations: application to breast MR images”. In: *IEEE transactions on medical imaging* 18.8 (1999), pp. 712–721.
- [17] W Paul Segars et al. “4D XCAT phantom for multimodality imaging research”. In: *Medical physics* 37.9 (2010), pp. 4902–4915.
- [18] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. “Deformable medical image registration: A survey”. In: *IEEE transactions on medical imaging* 32.7 (2013), pp. 1153–1190.
- [19] Andrew J Taylor et al. “T1 mapping: basic techniques and clinical applications”. In: *JACC: Cardiovascular Imaging* 9.1 (2016), pp. 67–81.
- [20] Nikolaus Weiskopf et al. “Quantitative magnetic resonance imaging of brain anatomy and in vivo histology”. In: *Nature Reviews Physics* 3.8 (2021), pp. 570–588.
- [21] Lukas Wissmann et al. “MRXCAT: Realistic numerical phantoms for cardiovascular magnetic resonance”. In: *Journal of Cardiovascular Magnetic Resonance* 16.1 (2014), pp. 1–11.
- [22] Zufeng Wu et al. “Medical image registration using b-spline transform”. In: *International Journal of Simulation: Systems, Science and Technology* 17.48 (2016), pp. 1–1.
- [23] Hui Xue et al. “Motion correction for myocardial T1 mapping using image registration with synthetic image estimation”. In: *Magnetic resonance in medicine* 67.6 (2012), pp. 1644–1655.
- [24] Yunlu Zhang et al. “GroupRegNet: a groupwise one-shot deep learning-based 4D image registration method”. In: *Physics in Medicine & Biology* 66.4 (2021), p. 045030.
- [25] Yongpei Zhu and Shi Lu. “Swin-voxelmorph: A symmetric unsupervised learning model for deformable medical image registration using swin transformer”. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VI*. Springer. 2022, pp. 78–87.